

Is it time for a Data Decadal Survey?

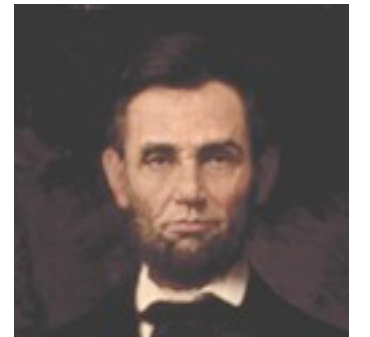
Anne Wilson, Ruth Duerr, Denise Hills, Sarah Ramdeen

ESIP 2013 Winter Meeting
January 10, 2013

Background

- Idea originated at Boulder Earth and Space Science Informatics Group (BESSIG) panel discussion on Science Policy for Engineers and Scientists, October 2012
- Suggested by Dr. Dan Baker, Director, Laboratory for Atmospheric and Space Physics (LASP)
- Authored “A Decadal Strategy for Solar and Space Physics,” released August 2012

The NAS



- National Academies
 - Provide advice to the nation on science, engineering, and medicine
 - Non-profit, private, not part of U.S. government
 - Comprised of four academies
 - National Academy of Science (NAS)
 - Created by law signed by Lincoln in 1863 to provide independent advice on scientific matters
 - 1916 established National Research Council (NRC)
 - 1964 established National Academy of Engineering (NAE)
 - 1970 established Institute of Medicine (IOM)

“Where the nation turns for Independent, Expert Advice”

The NRC

- “Operating arm” of NAS to conduct studies requested by and paid for by federal government or other sponsors.

“Our mission is to improve government decision making and public policy, increase public understanding, and promote the acquisition and dissemination of knowledge in matters involving science, engineering, technology, and health. The Research Council's independent, expert reports and other scientific activities inform policies and actions that have the power to improve the lives of people in the U.S. and around the world.”

- Conducts studies, surveys, writes reports
- Organized by Divisions, Boards
- Reports organized by Collections

NRC Divisions

- Earth and Life Studies
- Behavioral and Social Sciences and Education
- Engineering and Physical Sciences
- Policy and Global Affairs
- Transportation Research Board

Division on Earth and Physical Sciences (DEPS) Boards

Air Force Studies Board (AFSB)

Director Terry Jagers

Deputy Director Mike Clarke

Topics: Aeronautics, Defense, Engineering Research & Applications

Board on Army Science and Technology (BAST)

Director Bruce Braun

Topics: Climate Change/Environment, Defense, Materials

Board on Energy and Environmental Systems (BEES)

Director James Zucchetto

Topics: Climate Change/Environment, Energy

Board on Global Science and Technology

Joint administered with The Policy & Global Affairs Division

Director William Berry

Board on Infrastructure and the Constructed Environment (BICE)

Acting Director Dennis Chamot

Topics: Climate Change/Environment, Energy, Engineering Research & Applications, Infrastructure

Board on Mathematical Sciences and Their Applications (BMSA)

Director Scott T. Weidman

Topics: Defense, Information Technology, Mathematical Sciences & Applications, National Security & Intelligence

Board on Physics and Astronomy (BPA)

Director James C. Lancaster

Topics: Defense, Engineering Research & Applications, Information Technology, Materials, Physics & Astronomy, Space Science & Technology

Computer Science and Telecommunications Board (CSTB)

Director Jon Eisenberg

Topics: Defense, Energy, Engineering Research & Applications, Information Technology, Mathematical Sciences & Applications, National Security & Intelligence

Laboratory Assessments Board (LAB)

Director Jim McGee

Topics: Aeronautics, Defense, Federal Laboratories

National Materials and Manufacturing Board (NMMB)

Acting Director Dennis Chamot


Topics: Defense, Engineering Research & Applications, Information Technology, Infrastructure, Manufacturing, Materials, National Security & Intelligence

Naval Studies Board (NSB)

Director Charles F. Draper

Topics: Defense, Information Technology, National Security & Intelligence

NRC Division on Policy and Global Affairs, Board on Research Data and Information (BRDI)



BOARD ON RESEARCH DATA AND INFORMATION

Policy and Global Affairs

THE NATIONAL ACADEMIES

HOME | BACKGROUND | ACTIVITIES & EVENTS | PUBLICATIONS

Search

QUICK LINKS

- MEMBERS
- SPONSORS
- OTHER ORGANIZATIONS
- STAFF
- NEWSLETTER: BRDI TOLD ME...

THE mission of the Board on Research Data and Information is to improve the stewardship, policy, and use of digital data and information for science and the broader society.

STATEMENT OF TASK

The Board on Research Data and Information (BRDI) maintains surveillance of the field and proposes initiatives that might be undertaken at the National Research Council (NRC), targeted at challenges of national and international significance of particular interest to the board's sponsors. The Board engages in planning, program development, and administrative oversight of projects launched under its auspices. Acting through the NRC, BRDI focuses on the following tasks within its primary mission areas:

1. Address emerging issues in the management, policy, and use of research data and information at the national and international levels.
2. Through studies and reports of the NRC, provide independent and objective advice, reviews of programs, and assessment of priorities concerning research data and information activities and interests of its sponsors.
3. Encourage and facilitate collaboration across disciplines, sectors, and nations with regard to common interests in research data and information activities.
4. Monitor, assess, and contribute to the development of U.S. government and research community positions on research data and information programs and policies.
5. Initiate or respond to requests for consensus studies, workshops, conferences, and other activities within the Board's mission, and provide oversight for the activities performed under the Board's auspices.
6. Broadly disseminate and communicate the results of the Board's activities to its stakeholders and to the general public.

Events

JUST RELEASED

[The Future of Scientific Knowledge Discovery in Open Networked Environments: Summary of a Workshop](#)

[The Case for International Sharing of Scientific Data: A Focus on Developing Countries](#)

[For Attribution -- Developing Data Attribution and Citation Practices and Standards: Summary of an International Workshop](#)

UPCOMING EVENTS

Fourth Meeting of the BRDI Digital Curation study committee
January 23-24, 2013, Washington, DC
(Closed attendance for study committee members and staff only)

8th BRDI Board Meeting
February 26-27, 2013, Washington, DC

RECENT EVENTS

CODATA 2012: Open Data and Information for a Changing Planet
October 28-31, 2012, Taipei

Seventh BRDI Board Meeting
August 29-30, 2012, Washington, DC

BRDI Symposium: Global Scientific Data Infrastructures
August 29, 2012, Washington, DC

Contact Us

Board on Research Data and Information
Policy and Global Affairs Division
The National Academies
Room 511
500 Fifth Street, NW
Washington, DC 20001
USA
Email: publin@nas.edu
Phone: (202) 334-1531
Fax: (202) 334-2231

[Follow @PaulFUhlir](#)

We are grateful for the following grants for support of the Board on Research Data and Information:

*Institute of Museum and Library Services, IMLS: LG-00-09-0195-09;
National Institute of Standards and Technology, SB1341-09-CN-0095;
National Institutes of Health, NIH: N01-OD-4-2139;
National Oceanic and Atmospheric Administration, NA10OAR4600131;
National Science Foundation, OCI-1040898, and the
U.S. Geological Survey, G10AP00161.*

Some BRDI Publications



The Socioeconomic Effects of Public Sector Information on Digital Networks (2009)



Scientific Data for Decision Making Toward Sustainable Development: Senegal River Basin Case Study: Summary of a Workshop (2003)



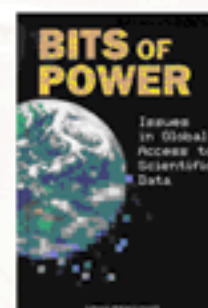
Strategies for Preservation of and Open Access to Scientific Data in China (2006)



A Question of Balance: Private Rights and the Public Interest in Scientific and Technical Databases (1999)



Open Access and the Public Domain in Digital Data and Information for Science: Proceedings of an International Symposium (2004)



Bits of Power: Issues in Global Access to Scientific Data (1997)



Electronic Scientific, Technical, and Medical Journal Publishing and Its Implications: Report of a Symposium (2004)



Preserving Scientific Data on Our Physical Universe (1995)



The Role of Scientific and Technical Data and Information in the Public Domain: Proceedings of a Symposium (2003)



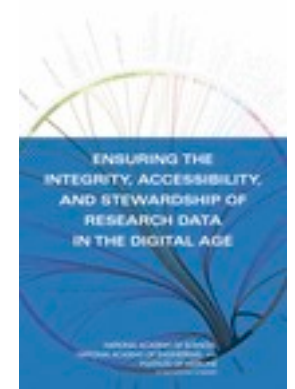
Finding the Forest in the Trees: The Challenge of Combining Diverse Environmental Data (1995)

NRC Reports

- Written by committees of expert volunteers from academia, industry, government, other organizations
- Most reports managed by one of more than 60 Boards
- “A Decadal Strategy for Solar and Space Physics”
 - requested by NASA and NSF
 - 300 white papers
 - town hall meetings, workshops
 - 85 NRC appointed participants
 - 18 member steering committee
 - \$200 - \$300K spent by NASA, NSF (I think...)
 - correction: costs much more -AW

2009 NRC report: “Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age”

“...examines the consequences of the changes affecting research data with respect to three issues - integrity, accessibility, and stewardship-and finds a need for a new approach to the design and the management of research projects.”



Under “Ethical Foundations of Science” Collection

Interoperability?

Discoverability?

Usability?

Report Recommendations (I)

Data Integrity Principle: Ensuring the integrity of research data is essential for advancing scientific, engineering, and medical knowledge and for maintaining public trust in the research enterprise. Although other stakeholders in the research enterprise have important roles to play, researchers themselves are ultimately responsible for ensuring the integrity of research data.

Recommendation 1: Researchers should design and manage their projects so as to ensure the integrity of research data, adhering to the professional standards that distinguish scientific, engineering, and medical research both as a whole and as their particular fields of specialization.

Recommendation 2: Research institutions should ensure that every researcher receives appropriate training in the responsible conduct of research, including the proper management of research data in general and within the researcher's field of specialization. Some research sponsors provide support for this training and for the development of training programs.

Recommendation 3: The research enterprise and its stakeholders—research institutions, research sponsors, professional societies, journals, and individual researchers—should develop and disseminate professional standards for ensuring the integrity of research data and for ensuring adherence to these standards. In areas where standards differ between fields, it is important that differences be clearly defined and explained. Specific guidelines for data management may require reexamination and updating as technologies and research practices evolve.

Recommendation 4: Research institutions, professional societies, and journals should ensure that the contributions of data professionals to research are appropriately recognized. In addition, research sponsors should acknowledge that financial support for data professionals is an appropriate component of research support in an increasing number of fields.

Report Recommendations (2)

Data Access and Sharing Principle: Research data, methods, and other information integral to publicly reported results should be publicly accessible.

Recommendation 5: All researchers should make research data, methods, and other information integral to their publicly reported results publicly accessible in a timely manner to allow verification of published findings and to enable other researchers to build on published results, except in unusual cases in which there are compelling reasons for not releasing data. In these cases, researchers should explain in a publicly accessible manner why the data are being withheld from release.

Recommendation 6: In research fields that currently lack standards for sharing research data, such standards should be developed through a process that involves researchers, research institutions, research sponsors, professional societies, journals, representatives of other research fields, and representatives of public interest organizations, as appropriate for each particular field.

Recommendation 7: Research institutions, research sponsors, professional societies, and journals should promote the sharing of research data through such means as publication policies, public recognition of outstanding data-sharing efforts, and funding.

Recommendation 8: Research institutions should establish clear policies regarding the management of and access to research data and ensure that these policies are communicated to researchers. Institutional policies should cover the mutual responsibilities of researchers and the institution in cases in which access to data is requested or demanded by outside organizations or individuals.

Report Recommendations (3)

Data Stewardship Principle: Research data should be retained to serve future uses. Data that may have long-term value should be documented, referenced, and indexed so that others can find and use them accurately and appropriately.

Recommendation 9: Researchers should establish data management plans at the beginning of each research project that include appropriate provisions for the stewardship of research data.

Recommendation 10: As part of the development of standards for the management of digital data, research fields should develop guidelines for assessing the data being produced in that field and establish criteria for researchers about which data should be retained.

Recommendation 11: Research institutions and research sponsors should study the needs for data stewardship by the researchers they employ and support. Working with researchers and data professionals, they should develop, support, and implement plans for meeting those needs.

Data is a national investment

- So is software

Data Issues

- Management of physical samples
- Loss of data due to technological changes
- Proliferation of data formats
- Data not machine readable
- Conflicting agency requirements
- Lack of understanding of a data set
- Dark data, lack of facilities for sharing
- Lack of knowledge of data sets available
- Badly managed software
- Lack of reproducibility
- Lack of unambiguous data set identification
- Lack of recognition for creation of high quality data sets
- Charged political climate, lack of trust in scientific results
- Lack of agreement of what a “data set” is (and how to spell the it!), collection, granules, ...
- etc., etc., etc.

These issues not addressed by 2009 report

Current efforts

- NASA: ESDSWG
 - NSF: EarthCube, DataOne, ...
 - AGU: ESSI
 - ESIP
 - NRC: RDA (international)
 - White House: data.gov
 - International Council for Science (ICSU): CODATA (international)
 - [your favorite effort here]
 - ...
-
- Agencies now requiring data management plans

Why a Decadal Survey?

- Cross agency/organization coherence
 - Currently ad hoc cross organizational efforts
- Data issues span domains
- Increasing cross domain science needs
- Demonstrated to be a robust method for developing goals and objectives
- Possibly faster achievement of identified goals
- Possibly better results?
- Cheaper overall than less coordinated individual agency efforts?

The Process

- NAS staff and board members work with sponsor to determine specific set of questions to be addressed
- Formal “statement of task” is developed that defines scope of study and serves as basis for determining expertise, balance of perspectives needed
- Sponsor funding options: contracts, grants, coop agreements, purchase orders
- May have one or several sponsors
- Typically takes 9 - 18 months
- NAS seeks nominees for committee members
- NAS president has sole authority for appointing all members
 - balance and conflict of interest discussion held in closed session during public comment period

For more than 140 years, the National Academies have been advising the nation on issues of science, technology, and medicine. An 1863 Congressional charter signed by President Lincoln authorized this non-governmental institution to honor top scientists with membership and to serve the nation whenever called upon. Today the National Academies—**National Academy of Sciences, National Academy of Engineering, Institute of Medicine, and National Research Council**—continue that dual mission.

Like no other organization, the National Academies can enlist the nation's foremost scientists, engineers, health professionals, and other experts to address the scientific and technical aspects of society's most pressing problems. Each year, more than 6,000 of these experts are selected to serve on hundreds of study committees that are convened to answer specific sets of questions. All serve without pay.

The National Academies are a unique resource. Their reports are viewed as being valuable and credible because of their reputation for providing advice with high standards of scientific and technical quality and independence. National Academies staff will work with potential sponsors to develop a specific set of questions to be answered by a committee of experts.

This guide is intended for prospective sponsors interested in requesting studies from the National Academies. It describes the process for producing these reports—from funding to report dissemination—and explains the sponsors' involvement at each stage. This approach ensures that sponsors receive the best product possible. Checks and balances are applied at every step in the study process to protect the integrity of the reports and to maintain public confidence in them.

Defining and Initiating the Study

Before a contract or grant is signed, National Academies staff and board members work with the sponsor to determine the specific set of questions to be addressed. A formal "statement of task" is developed that defines the scope of the study and serves as the basis for determining the expertise and the balance of perspectives needed on the committee.

Typically, the following steps are taken to ensure that the project is clearly defined and that both the National Academies and the sponsor understand what is expected throughout the study process:

1. Sponsor and National Academies staff meet to discuss task, schedule, and likely costs.
2. When the staff and sponsor have come to a general agreement, staff prepares a prospectus for approval by the Executive Committee of the National Research Council Governing Board. This step ensures the appropriateness of the topic and the scope of the study before a formal proposal is sent to the sponsor.

Most studies are funded by those requesting the advice. Consistent with the congressional charter, experts serving on study committees volunteer their time without compensation. The cost of consensus studies can range from about \$200,000 to more than \$1 million, depending on the breadth and complexity of the issues being addressed and the

length of time needed to produce the desired report. The costs include the expenses of committee meetings, professional staff supporting the committee, report publication, and public dissemination.

Several funding options are available to sponsors. These include contracts, grants, cooperative agreements, and purchase orders. A project may have one or several sponsors. Studies typically take from 9 to 18 months to complete, although some urgent studies may be completed in a shorter time and some broader or more complex studies may take a longer time.

Because of their unique origins and status, federal sponsors interested in having the National Academies conduct a study can obtain their services on a sole source basis. The National Academies do not compete for federal contracts. Grants, contracts, and gifts can also be received from states, foundations, and other sources.

A project begins after formal commitments have been received for sufficient funds to complete the study.

Committee Selection and Approval

Selection of appropriate committee members, individually and collectively, is essential for the success of every study. All committee members serve as individual experts, not as representatives of organizations or interest groups. Each member is expected to contribute to the project on the basis of his or her own expertise and good judgment.

Once a project is funded, the National Academies seek nominees for members of consensus study committees from many sources, including the sponsors. However, the president of the National Academy of Sciences (who also serves as chair of the National Research Council) has the sole authority for appointing all members of study committees. Before a committee can be approved, a thorough balance and conflict of interest discussion is held in closed session during the first meeting, and any issues raised in that discussion or during the public comment period are investigated and addressed. Committee members are considered prospective until after this process is completed.

Committee Meetings, Deliberations, and Drafting Report

Study committees typically gather information through: 1) meetings that are open to the public and announced in advance through the National Academies' Web site, 2) submission of information by outside parties, 3) reviews of scientific literature, and 4) investigations by the committee members and staff. In all cases, efforts are made to solicit input from individuals who have been directly involved in, or who have special knowledge of, the topic under consideration. Sponsors are typically invited to make presentations to the committee at its first couple of meetings to discuss the sponsors' expectations for the study. Also, the sponsor is asked to provide as much information relevant to the study as possible.

SPONSOR INTERACTIONS

Sponsor and Academies staff meet to discuss scope, timing and likely cost of project

Sponsor explores available funding options

Staff seek nominees for committee members from sponsor and other sources

Committee's first meeting (sponsor invited to participate in open session)

Public information gathering sessions (sponsors invited)

Sponsor provided with copy of final report and offered a briefing

DEFINING AND INITIATING THE STUDY

COMMITTEE SELECTION AND APPROVAL

REPORT REVIEW

REPORT RELEASE AND DISSEMINATION

NRC Governing Board reviews and approves project

Project starts when funding is formally approved

Prospective membership posted at least 20 days for public comment

COMMITTEE MEETINGS, DELIBERATIONS, AND DRAFTING REPORT

Deliberation and report drafting (closed session)

Full committee signs off on draft report

Committee and Academies sign off on report

Report released to the public

STUDY MILESTONES

Possible Outcomes

- Coordinated cross agency vision for data management, interoperability and usability
- Identification of highest priority needs
- Path for moving forward, with resource limitations in mind
- Identification of metrics for success
- Recommendations
 - Funding priorities
 - Ways to provide practitioner support
 - A cross agency body to help practitioners?
- Resolve conflicting requirements from agencies
- Faster elimination of cultural barriers

Questions, Issues

- Are any agencies interested enough in having such a report that they would be willing to sponsor it?
- Which NRC Board(s) would perform the survey?
- Scope: where would the line be drawn? How to determine?
- Software is also very important!
- Earth sciences? But would like to enable broader cross domain research...
- Break problem into multiple areas?

What's next?

- Determine focus
 - Goal: List, group, prioritize issues that are worth tackling
- Continue to contact agency representatives, other possible sponsors, seek input as to their areas of concern
- Form a cluster?